PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C07H 21/04, C07K 1/00, C12N 5/00 | A1

(11) International Publication Number: WO 00/12532

(43) International Publication Date: 9 March 2000 (09.03.00)

(21) International Application Number: PCT/US99/19047

(22) International Filing Date: 20 August 1999 (20.08.99)

(30) Priority Data:
60/098,521 31 August 1998 (31.08.98) US

(71) Applicant (for all designated States except US): ELI LILLY AND COMPANY [US/US]; Lilly Corporate Center, Drop Code 1104, Indianapolis, IN 46285 (US).

(72) Inventors; and
(75) Inventors/Applicants (for US only): ROSTECK, Paul, Robert, Junior [US/US]; 4247 Saffron Drive, Indianapolis, IN 46237 (US). SU, Wen [US/US]; 13447 Dunes Drive, Carmel, IN 46032 (US). LI, Xuan, Mei [CN/US]; 3811 Brigade Court, Carmel, IN 46032 (US).

(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
With international search report.

(54) Title: FACTOR 8 HOMOLOGUE NUCLEIC ACIDS, POLYPEPTIDES, METHODS, USES

(57) Abstract

The present invention relates to at least one novel F8H polypeptide, including isolated nucleic acids that encode at least one F8H polypeptide, F8H polypeptides, vectors, host cells, trangenics, chimerics, and methods of making and using thereof same, as well as F8H-specific antibodies and methods.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania
AM	Armenia
AT	Austria
AU	Australia
AZ	Azerbaijan
BA	Bosnia and Herzegovina
BB	Barbados
BE	Belgium
BF	Burkina Faso
BG	Bulgaria
BJ	Benin
BR	Brazil
BY	Belarus
CA	Canada
CF	Central African Republic
CG	Congo
CH	Switzerland
CI	Côte d'Ivoire
CM	Cameroon
CN	China
CU	Cuba
CZ	Czech Republic
DE	Germany
DK	Denmark
EE	Estonia
ES	Spain
FI	Finland
FR	France
GA	Gabon
GB	United Kingdom
GE	Georgia
GH	Ghana
GN	Guinea
GR	Greece
HU	Hungary
IE	Ireland
IL	Israel
IS	Iceland
IT	Italy
JP	Japan
KE	Kenya
KG	Kyrgyzstan
KP	Democratic People's Republic of Korea
KR	Republic of Korea
KZ	Kazakstan
LC	Saint Lucia
LI	Liechtenstein
LK	Sri Lanka
LR	Liberia
LS	Lesotho
LT	Lithuania
LU	Luxembourg
LV	Latvia
MC	Monaco
MD	Republic of Moldova
MG	Madagascar
MK	The former Yugoslav Republic of Macedonia
ML	Mali
MN	Mongolia
MR	Mauritania
MW	Malawi
MX	Mexico
NE	Niger
NL	Netherlands
NO	Norway
NZ	New Zealand
PL	Poland
PT	Portugal
RO	Romania
RU	Russian Federation
SD	Sudan
SE	Sweden
SG	Singapore
SI	Slovenia
SK	Slovakia
SN	Senegal
SZ	Swaziland
TD	Chad
TG	Togo
TJ	Tajikistan
TM	Turkmenistan
TR	Turkey
TT	Trinidad and Tobago
UA	Ukraine
UG	Uganda
US	United States of America
UZ	Uzbekistan
VN	Viet Nam
YU	Yugoslavia
ZW	Zimbabwe

FACTOR 8 HOMOLOGUE NUCLEIC ACIDS, POLYPEPTIDES, METHODS, USES

CROSS REFERENCE

This application claims the benefit of U.S. Provisional Application No. 60/098,521, filed August 31, 1998, which application is entirely incorporated herein by reference.

BACKGROUND OF THE INVENTION

FIELD OF THE INVENTION

The present invention relates to compounds and compositions comprising novel Factor 8 Homologue (F8H) polypeptides and nucleic acids. More specifically, recombinant or isolated nucleic acid molecules are provided encoding human F8H polypeptides as well as vectors, host cells, antibodies and methods for producing and using different aspects of the invention.

RELATED ART

Factor 8 is a plasma protein which plays an essential role in the coagulation system. When assembled with Factor IXa on a phospholipid membrane, it functions as a cofactor in an enzyme complex that cleaves the zymogen Factor X to Factor Xa.

Mutations that inactivate Factor 8 lead to hemophilia in humans. Inversion mutations of Factor 8 account for about 45% of severe hemophilia A cases.

Factor 8 and Factor 5 are homologous coagulation cofactors sharing a similar domain organization (A1-A2-B-A3-C1-C2), and both are extensively glycosylated within their B-domains.

The Factor 8 homologue described herein contains a Factor 5/8 signature, and is useful as a therapeutic protein for treating coagulation related diseases, such as hemophilia and stroke. The gene product of the present invention is also useful as a target to screen for modulator(s) of the Factor 8 homologue, said modulator(s) being useful for treating diseases such as hemophilia and stroke.

SUMMARY OF THE INVENTION

The present invention provides isolated F8H nucleic acids and encoded polypeptides, including fragments and specified variants thereof, as well as F8H compositions, probes, vectors, host cells, antibodies, and methods, as described and enabled herein.

The present invention provides, in one aspect, isolated nucleic acid molecules comprising a polynucleotide encoding specific F8H polypeptides, as fragments or specified variants, comprising at least one domain thereof.

For example, such polypeptides are provided as fragments and/or variants corresponding to at least five contiguous amino acid residues of SEQ ID NO:2.

The present invention further provides recombinant vectors, comprising said isolated F8H nucleic acid molecules of the present invention, host cells containing such nucleic acids and/or recombinant vectors, as well as methods of making and/or using such nucleic acid, vectors and/or host cells.

The present invention also provides methods of making or using such nucleic acids, vectors and/or host cells, such

as, but not limited to, use in the production of F8H nucleic acids and/or polypeptides by known recombinant, synthetic and/or purification techniques, based on the teaching and guidance presented herein in combination with what is known in the art.

The present invention also provides an isolated F8H polypeptide, comprising at least one fragment, domain or specified variant of at least 5 contiguous amino acids of SEQ ID NO:2.

The present invention also provides an isolated F8H polypeptide comprising SEQ ID NO:2.

The present invention also provides an isolated F8H polypeptide, as described herein, wherein the polypeptide further comprises at least one specified mutation corresponding to at least one residue of SEQ ID NO:2.

The present invention also provides an isolated F8H polypeptide, as described herein, wherein the polypeptide has activity as a coagulation factor.

The present invention also provides a composition comprising an isolated F8H nucleic acid and/or polypeptide as described herein, and a carrier or diluent. The carrier or diluent can optionally be pharmaceutically acceptable, according to known methods.

The present invention also provides an isolated nucleic acid probe, wherein the nucleic acid comprises a polynucleotide of at least 10 nucleotides, corresponding or complementary to at least 10 nucleotides of at least one of the sequences disclosed herein.

The present invention also provides a vector comprising an isolated F8H nucleic acid as described herein.

The present invention also provides a vector, as described herein, wherein the vector is linear or circular, single or double stranded, DNA, RNA, or a combination thereof.

The present invention also provides a host cell, comprising an isolated F8H nucleic acid as described herein.

The present invention also provides a method for constructing a recombinant host cell that expresses an F8H polypeptide, comprising introducing into the host cell an F8H nucleic acid in replicatable form. The present invention also provides a recombinant host cell provided by a method described herein.

The present invention also provides a method for expressing an F8H polypeptide in a recombinant host cell, comprising culturing a recombinant host cell as described herein under conditions wherein the F8H polypeptide is expressed in detectable or recoverable amounts.

The present invention also provides an isolated F8H polypeptide produced by a recombinant, synthetic, and/or any suitable purification method as described herein and/or as known in the art.

The present invention also provides an F8H antibody or fragment, comprising a polyclonal and/or monoclonal antibody or fragment that specifically binds at least one epitope specific to an F8H polypeptide as described herein.

The present invention also provides a method for producing an F8H antibody or antibody fragment thereof,

comprising generating the antibody or fragment by recombinant, synthetic and/or hybridoma techniques.

The present invention also provides an F8H antibody or fragment produced by a method as described herein.

The present invention also provides a method for identifying compounds that bind an F8H polypeptide, comprising

a) admixing at least one isolated F8H polypeptide as described herein with a test compound or composition; and

b) detecting at least one binding interaction between the polypeptide and the compound or composition, optionally further comprising detecting a change in the biological activity of said F8H polypeptide.

DESCRIPTION OF THE INVENTION

The present invention provides isolated, recombinant and/or synthetic nucleic acid molecules encoding F8H polypeptides, specific fragments and specified variants thereof, and methods of making and using said nucleic acids and polypeptides.

Utility

The present invention provides at least one utility by providing isolated F8H nucleic acids of sufficient length and complementarity to an F8H nucleic acid for use as a hybridization probe or amplification primer in the detection, quantitation, or isolation of gene sequences or transcripts identical to or related to F8H. For example, isolated nucleic acids of the present invention can be used as probes in detecting deficiencies in the level of F8H mRNA, in

screens for the detection of mutations in the F8H gene homologue (e.g., substitutions, deletions, or additions), or for monitoring regulation of expression of said gene.

The isolated nucleic acids of the present invention can also be used for recombinant expression of F8H polypeptides, or for use as immunogens in the preparation and/or screening of antibodies. The isolated nucleic acids of the present invention can also be employed for use in sense or antisense suppression of one or more F8H genes or nucleic acids, in a host cell, or tissue, in vivo or in vitro. Attachment of chemical agents which bind, intercalate, cleave and/or crosslink to the isolated nucleic acids of the present invention can also be used to modulate transcription or translation of at least one nucleic acid disclosed herein.

Citations

All references to publications cited herein are entirely incorporated herein by reference as they show the state of the art at the time of the present invention to provide description and enablement of the present invention. Publications refer to scientific, patent publication or any other information available in any media format, including all recorded, electronic or printed formats. The following citations are entirely incorporated by reference: Ausubel, *et al.*, ed., *Current Protocols in Molecular Biology*, Greene Publishing, NY, NY (1987-1988); Sambrook, *et al.*, *Molecular Cloning: A Laboratory Manual*, 2nd Edition, Cold Spring Harbor, NY (1989); Harlow and Lane, *Antibodies, a Laboratory Manual*, Cold Spring Harbor, NY (1989); Colligan, *et al.*,

eds., *Current Protocols in Immunology*, Greene Publishing, NY (1994-1988).

Definitions

The terms "complementary" or "complementarity" as used herein refer to the capacity of purine and pyrimidine nucleotides to associate through hydrogen bonding to form double stranded nucleic acid molecules. The following base pairs are related by complementarity: guanine and cytosine; adenine and thymine; and adenine and uracil. As used herein, "complementary" means that the aforementioned relationship applies to substantially all base pairs comprising two single-stranded nucleic acid molecules over the entire length of said molecules. "Partially complementary" refers to the aforementioned relationship in which one of two single-stranded nucleic acid molecules is shorter in length than the other such that a portion of one of the molecules remains single-stranded.

The term "conservative modification" or "conservative substitution" refers to a change in one or a plurality of base pairs in a nucleic acid disclosed herein, or to one or a plurality of amino acid residues in a protein disclosed herein. Such modifications may, for example, comply with Table 1.

Table 1:

ORIGINAL RESIDUE	EXEMPLARY SUBSTITUTIONS
ALA	SER
ARG	LYS
ASN	GLN; HIS
ASP	GLU
CYS	SER
GLN	ASN
GLU	ASP
GLY	PRO
HIS	ASN, GLN
ILE	LEU, VAL
LEU	ILE, VAL
LYS	ARG, GLN, GLU
MET	LEU, ILE
PHE	MET, LEU, TYR
SER	THR
THR	SER
TRP	TYR
TYR	TRP, PHE
VAL	ILE, LEU

"Expressed sequence tag" (viz. EST) refers to a fragment of a full-length cDNA sequence. EST fragments can encode active or inactive peptide fragments and are useful, among other things, as nucleic hybridization probes.

"Fragment thereof" refers to a fragment, piece, or sub-region of a nucleic acid or protein molecule whose sequence is disclosed herein, such that said fragment comprises 5 or

more amino acids, or 10 or more nucleotides that are contiguous in the parent protein or nucleic acid molecule.

"Functional fragment" as used herein, refers to an isolated sub-region, or fragment of a protein disclosed herein, or sequence of amino acids that, for example, comprises a functionally distinct region such as an active site for an enzyme, or a binding site for a substrate, or a binding site for a receptor. Functional fragments may be produced by cloning technology, or as the natural products of alternative splicing mechanisms.

"Functionally related" as used herein is applied to proteins or peptides that are predicted to be functionally similar or identical to a progenitor molecule, for example, Dhh-H or fragment thereof. Functionally related species are identified based on chemical and physical similarities in amino acid composition and sequence.

"Host cell" refers to any eucaryotic or procaryotic cell that is suitable for propagating and/or expressing a cloned gene contained on a vector that is introduced into said host cell by, for example, transformation or transfection, or the like.

"F8H" refers to a gene or RNA (e.g. SEQ ID NO:1 or the RNA equivalent thereof; viz. "U" replaces "T") and a protein (SEQ ID NO:2). F8H is a member of the family of blood coagulation proteins.

The term "homolog" or "homologous" describes the relationship between different nucleic acid molecules or amino acid sequences in which said sequences or molecules

are related by partial identity or similarity at one or more blocks or regions within said molecules or sequences.

The term "hybridization" as used herein refers to a process in which a single-stranded nucleic acid molecule joins with a complementary strand through nucleotide base pairing. "Selective hybridization" refers to hybridization under conditions of high stringency. The degree of hybridization depends upon, for example, the degree of homology, the stringency of hybridization, and the length of hybridizing strands.

"Isolated nucleic acid compound" refers to any RNA or DNA sequence, however constructed or synthesized, which is locationally distinct from its natural location.

A "nucleic acid probe" or "probe" as used herein is a labeled nucleic acid compound which hybridizes with another nucleic acid compound. "Nucleic acid probe" means a single stranded nucleic acid sequence that will combine with a complementary or partially complementary single stranded target nucleic acid sequence to form a double-stranded molecule. A nucleic acid probe may be an oligonucleotide or a nucleotide polymer. A probe will usually contain a detectable moiety which may be attached to the end(s) of the probe or be internal to the sequence of the probe.

The term "orthologue" or "orthologous" refers to two or more genes or proteins from different organisms that exhibit sequence homology.

The term "paralogue" or "paralogous" refers to two or more genes or proteins within a single organism that exhibit sequence homology.

The term "plasmid" refers to an extrachromosomal genetic element. The plasmids disclosed herein are commercially available, publicly available on an unrestricted basis, or can be constructed from readily available plasmids in accordance with published procedures.

A "primer" is a nucleic acid fragment which functions as an initiating substrate for enzymatic or synthetic elongation of, for example, a nucleic acid molecule.

The term "promoter" refers to a nucleic acid sequence that directs transcription, for example, of DNA to RNA. An inducible promoter is one that is regulatable by environmental signals, such as carbon source, heat, or metal ions, for example. A constitutive promoter generally operates at a constant level and is not regulatable.

The terms "protein" and "polypeptide" are used interchangeably herein as a biopolymer comprising at least 10 amino acid residues.

"Recombinant DNA cloning vector" as used herein refers to any autonomously replicating agent, including, but not limited to, plasmids and phages, comprising a DNA molecule to which one or more additional DNA segments can or have been incorporated.

The term "recombinant DNA expression vector" or "expression vector" as used herein refers to any recombinant DNA cloning vector, for example a plasmid or phage, in which a promoter and other regulatory elements are present thereby enabling transcription of an inserted DNA, which may encode a protein.

The term "stringency" refers to hybridization conditions. High stringency conditions disfavor non-homologous base pairing. Low stringency conditions have the opposite effect. Stringency may be altered, for example, by temperature and salt concentration.

"Low stringency" conditions comprise, for example, a temperature of about 37° C or less, a formamide concentration of less than about 50%, and a moderate to low salt (SSC) concentration; or, alternatively, a temperature of about 50° C or less, and a moderate to high salt (SSPE) concentration, for example 1M NaCl.

"High stringency" conditions comprise, for example, a temperature of about 42° C or less, a formamide concentration of less than about 20%, and a low salt (SSC) concentration; or, alternatively, a temperature of about 65° C, or less, and a low salt (SSPE) concentration. For example, high stringency conditions comprise hybridization in 0.5 M $NaHPO_4$, 7% sodium dodecyl sulfate (SDS), 1 mM EDTA at 65°C (Ausubel, F.M. *et al.* Current Protocols in Molecular Biology, Vol. I, 1989; Green Inc. New York, at 2.10.3).

"SSC" comprises a hybridization and wash solution. A stock 20X SSC solution contains 3M sodium chloride, 0.3M sodium citrate, pH 7.0.

"SSPE" comprises a hybridization and wash solution. A 1X SSPE solution contains 180 mM NaCl, 9mM Na2HPO4, 0.9 mM NaH2PO4 and 1 mM EDTA, pH 7.4.

"Substantially pure," used in reference to a peptide or protein, means separation from other cellular and non-

cellular molecules, including other protein molecules. A substantially pure preparation would be about at least 85% pure; preferably about at least 95% pure. A "substantially pure" protein can be prepared by a variety of techniques, well known to the skilled artisan, including, for example, the IMAC protein purification method.

"Treating" as used herein describes the management and care of a patient for the purpose of combating the disease, condition, or disorder and includes the administration of a protein of the present invention to prevent the onset of the symptoms or complications, alleviating the symptoms or complications, or eliminating the disease, condition, or disorder. Treating as used herein includes the administration of the protein for cosmetic purposes. A cosmetic purpose seeks to control, for example, the weight of a mammal to improve bodily appearance.

The term "variant" or "specified variant" as used herein means a nucleic acid or protein that is structurally and/or functionally similar to a sequence disclosed herein, but not identical.

The term "vector" as used herein refers to a nucleic acid compound used for introducing exogenous or endogenous DNA into host cells. A vector comprises a nucleotide sequence which may encode one or more protein molecules. Plasmids, cosmids, viruses, and bacteriophages, in the natural state or which have undergone recombinant engineering, are examples of commonly used vectors.

The various restriction enzymes disclosed and described herein are commercially available and the manner of use of

said enzymes including reaction conditions, cofactors, and other requirements for activity are well known to one of ordinary skill in the art. Reaction conditions for particular enzymes were carried out according to the manufacturer's recommendation.

Nucleic Acid Molecules

Unless otherwise indicated, all nucleotide sequences identified by sequencing a nucleic acid molecule herein were identified using an automated nucleic acid sequencer, and all amino acid sequences of polypeptides encoded by nucleic acid molecules identified herein were identified by codon correspondence, or by translation of a nucleic acid sequence identified herein.

Using the information provided herein, such as the nucleotide sequences encoding at least a 3-50 amino acid fragment of SEQ ID NO:2 or a deposited vector comprising SEQ ID NO:1, a nucleic acid molecule of the present invention encoding an F8H polypeptide can be obtained.

The identified nucleotide sequence of an F8H nucleic acid of SEQ ID NO:1 contains an open reading frame encoding a polypeptide identified herein by SEQ ID NO:2

As indicated, nucleic acid molecules of the present invention can be in the form of RNA, such as mRNA, or hnRNA, or in the form of DNA, including, but not limited to, cDNA and genomic DNA obtained by cloning or produced synthetically, or any combination thereof. The DNA can be triple-stranded, double-stranded or single-stranded, or any combination thereof.

Isolated nucleic acid molecules of the present invention include nucleic acid molecules comprising an open reading frame (ORF) shown in at least SEQ ID NO:1; nucleic acid molecules comprising the coding sequence for an F8H polypeptide; and nucleic acid molecules which comprise a nucleotide sequence substantially different from those described above but which, due to the degeneracy of the genetic code, still encode at least one F8H polypeptide as described and enabled herein. Thus, it would be routine for one skilled in the art to generate such degenerate nucleic acid variants that code for specific F8H polypeptides of the present invention. *See, e.g.,* Ausubel, *et al.*

In a further embodiment, nucleic acid molecules are provided encoding the mature F8H polypeptide or the full-length F8H polypeptide lacking the N-terminal methionine. The invention also provides an isolated nucleic acid molecule having the nucleotide sequence designated herein as SEQ ID NO:1, or a nucleic acid molecule having a sequence complementary thereto.

Nucleic Acid Fragments

The present invention is further directed to fragments of the isolated nucleic acid molecules described herein. By a fragment of an isolated nucleic acid molecule is meant a molecule having at least 10 nucleotides of a nucleotide sequence of a deposited cDNA, or a nucleotide sequence shown in at least SEQ ID NO:1, and is intended to mean fragments at least about 10 nucleotides, at least about 15 nucleotides, at least about 30 nucleotides, and at least about 40 nucleotides in length, which are useful, *inter alia*

as diagnostic probes and primers as described herein. Of course, larger fragments such as at least about 50, 100, 120, 200, 500, 1000, 1500, 2000, 2500, 3000, 3500, and/or 4000 nucleotides in length, are also useful according to the present invention. By a fragment at least 10 nucleotides in length, for example, is intended fragments which include 10 or more contiguous bases from the nucleotide sequence of SEQ ID NO:1.

Such nucleotide fragments are useful according to the present invention for screening DNA sequences that code for one or more fragments of an F8H polypeptide as described herein. Such screening, as a non-limiting example, can include the use of so-called "DNA chips" for screening DNA sequences of the present invention of varying lengths, as described, e.g., in U.S. Patent Nos. 5,631,734, 5,624,711, 5,744,305, 5,770,456, 5,770,722, 5,675,443, 5,695,940, 5,710,000, 5,733,729, which are entirely incorporated herein by reference.

As indicated, nucleic acid molecules of the present invention can include, but are not limited to, those encoding the F8H amino acid sequence of the mature polypeptide; the coding sequence of the mature polypeptide and additional sequences, such as non-coding sequences, including, for example, but not limited to introns and non-coding 5' and 3' sequences, such as the transcribed, non-translated sequences that play a role in transcription, mRNA processing, including splicing and polyadenylation signals (for example - ribosome binding and stability of mRNA); an additional coding sequence which codes for additional amino

acids, such as those which provide additional functionalities. Thus, the sequence encoding a polypeptide can be fused to a marker sequence, such as a sequence encoding a peptide that facilitates purification of the fused polypeptide.

Preferred nucleic acid fragments of the present invention also include nucleic acid molecules encoding epitope-bearing portions of an F8H polypeptide.

Oligonucleotide and Polynucleotide Probes

In another aspect, the invention provides a polynucleotide (either DNA or RNA) that comprises at least about 10 nucleotides ("nt"), and more preferably at least about 20 nt, still more preferably at least about 30 nucleotides, and even more preferably at least about 30-2000 nucleotides of a nucleic acid molecule described herein. These are useful as diagnostic probes and primers as discussed above and in more detail below.

By a portion of a polynucleotide of "at least 10 nucleotides in length," for example, is intended 10 or more contiguous nucleotides from the nucleotide sequence of the reference polynucleotide (e.g., at least one deposited nucleic acid or at least one nucleotide sequence as shown in SEQ ID NO:1.

Of course, a polynucleotide which hybridizes only to a poly A sequence (such as the 3' terminal poly(A) of an F8H cDNA shown in SEQ ID NO:1 or to a complementary stretch of T (or U) resides, would not be included in a probe of the invention, since such a polynucleotide would hybridize to any nucleic acid molecule containing a poly (A) stretch or

the complement thereof (e.g., practically any double-stranded cDNA clone).

The present invention also provides subsequences of full-length nucleic acids. Any number of subsequences can be obtained by reference to SEQ ID NO:1 using primers which selectively amplify, under stringent conditions to at least two sites to the polynucleotides of the present invention, or to two sites within the nucleic acid which flank and comprise a polynucleotide of the present invention, or to a site within a polynucleotide of the present invention and a site within the nucleic acid which comprises it. A variety of methods for obtaining 5' and/or 3' ends is well known in the art. See, e.g., RACE (Rapid Amplification of Complementary Ends) as described in M. A. Frohman, *PCR Protocols: A Guide to Methods and Applications*, M. A. Innis, D. H. Gelfand, J. J. Sninsky, T. J. White, Eds., Academic Press, Inc., San Diego, CA, pp. 28-38 (1990); see also, U.S. Patent No. 5,470,722, and Ausubel, *et al.*, *Current Protocols in Molecular Biology, Unit 15.6*, Eds., Greene Publishing and Wiley-Interscience, New York (1989-1998). Thus, the present invention provides F8H polynucleotides having the sequence of the F8H gene, nuclear transcript, cDNA, or complementary sequences and/or subsequences thereof.

Primer sequences can be obtained by reference to a contiguous subsequence of a polynucleotide of the present invention. Primers are chosen to selectively hybridize, under PCR amplification conditions, to a polynucleotide of the present invention in an amplification mixture comprising a genomic and/or cDNA library from the same species. In some

embodiments, the primers will be constructed to anneal at their 5′ terminal ends to the codon encoding the carboxy or amino terminal amino acid residue (or the complements thereof) of the polynucleotides of the present invention. The primer length in nucleotides is selected from the group of integers consisting of from at least 15 to 50. Thus, the primers can be at least 15, 18, 20, 25, 30, 40, or 50 nucleotides in length. A non-annealing sequence at the 5′ end of the primer (a "tail") can be added, for example, to introduce a cloning site at the terminal ends of the amplified DNA.

The amplification primers may optionally be elongated in the 3′ direction with additional contiguous nucleotides from the polynucleotide sequences. The number of nucleotides by which the primers can be elongated is selected from the group of integers consisting of from at least 1 to at least 25. Thus, for example, the primers can be elongated with an additional 1, 5, 10, or 15 nucleotides. Those of skill will recognize that a lengthened primer sequence can be employed to increase specificity of binding (i.e., annealing) to a target sequence.

The amplification products can be translated using expression systems well known to those of skill in the art and as discussed, infra. The resulting translation products can be confirmed as polypeptides of the present invention by, for example, assaying for the appropriate catalytic activity (e.g., specific activity and/or substrate specificity), or verifying the presence of one or more linear epitopes which are specific to a polypeptide of the present invention.

Methods for protein synthesis from PCR derived templates are known in the art and available commercially. See, e.g., Amersham Life Sciences, Inc., Catalog '97, p. 354.

Polynucleotides Which Selectively Hybridize to a Polynucleotide as Described Herein

The present invention provides isolated nucleic acids that hybridize under selective hybridization conditions to a polynucleotide disclosed herein, e.g., SEQ ID NO:1.

Screening polypeptides for specific binding to antibodies can be conveniently achieved using peptide display libraries. This method screens large collections of peptides for individual members having the desired function or structure. Antibody screening of peptide display libraries is well known in the art. The displayed peptide sequences can be from 3 to 5000 or more amino acids in length, frequently from 5-100 amino acids long, and often from about 8 to 15 amino acids long.

In addition to direct chemical synthetic methods for generating peptide libraries, several recombinant DNA methods have been described. One type involves the display of a peptide sequence on the surface of a bacteriophage or cell. Each bacteriophage or cell contains the nucleotide sequence encoding the particular displayed peptide sequence. Such methods are described in PCT Patent Publication Nos. 91/17271, 91/18980, 91/19818, and 93/08278. Other systems for generating libraries of peptides have aspects of both in vitro chemical synthesis and recombinant methods. See, PCT Patent Publication Nos. 92/05258, 92/14843, and 96/19256. See also, U.S. Patent Nos. 5,658,754; and 5,643,768. Peptide

display libraries, vector, and screening kits are commercially available from such suppliers as Invitrogen (Carlsbad, CA).

Polynucleotides Complementary to the Polynucleotides

As indicated above, the present invention provides isolated nucleic acids comprising F8H polynucleotides, wherein the polynucleotides are complementary to the polynucleotides described herein. Fully complementary sequences base-pair throughout the entirety of their length with such polynucleotides.

Construction of Nucleic Acids

The isolated nucleic acids of the present invention can be made using standard recombinant methods, synthetic techniques, purification techniques, or combinations thereof, as is well known in the art.

The nucleic acids may conveniently comprise sequences in addition to a polynucleotide of the present invention. For example, a multi-cloning site comprising one or more endonuclease restriction sites may be inserted into the nucleic acid to aid in isolation of the polynucleotide. Also, translatable sequences may be inserted to aid in the isolation of the translated polypeptide of the present invention. For example, a short histidine peptide may be added as a convenient means to purify the proteins of the present invention. The nucleic acid of the present invention - excluding the polynucleotide sequence - is optionally a vector, adapter, or linker for cloning and/or expression of a polynucleotide of the present invention.

Additional sequences may be added to such cloning and/or expression sequences to optimize their function in cloning and/or expression, to aid in isolation of the polynucleotide, or to improve the introduction of the polynucleotide into a cell. Typically, the length of a nucleic acid of the present invention less the length of its polynucleotide of the present invention is less than 20 kilobase pairs ("kb"), often less than 15 kb, and frequently less than 10 kb. Use of cloning vectors, expression vectors, adapters, and linkers is well known in the art.

Recombinant Methods for Constructing Nucleic Acids

The isolated nucleic acid compositions of this invention, such as RNA, cDNA, genomic DNA, or a hybrid thereof, can be obtained from biological sources using any number of cloning methodologies known to those of skill in the art. In some embodiments, oligonucleotide probes which selectively hybridize, under stringent conditions, to the polynucleotides of the present invention are used to identify the desired sequence in a cDNA or genomic DNA library. While isolation of RNA, and construction of cDNA and genomic libraries is well known to those of ordinary skill in the art.

Nucleic Acid Screening and Isolation Methods

A cDNA or genomic library can be screened using a probe based upon the sequence of a polynucleotide of the present invention, such as those disclosed herein. Probes may be used to hybridize with genomic DNA or cDNA sequences to isolate homologous genes in the same or different organisms.

Those of skill in the art will appreciate that various degrees of stringency of hybridization can be employed in the assay. The degree of stringency can be controlled by temperature, ionic strength, pH and the presence of a partially denaturing solvent such as formamide. For example, the stringency of hybridization is conveniently varied by changing the polarity of the reactant solution through, for example, manipulation of the concentration of formamide within the range of 0% to 50%. The degree of complementarity (sequence identity) required for detectable binding will vary in accordance with the stringency of the hybridization medium and/or wash medium. The degree of complementarity will optimally be 100%; however, it should be understood that minor sequence variations in the probes and primers may be compensated for by reducing the stringency of the hybridization and/or wash medium.

Methods of amplification of RNA or DNA are well known in the art and can be used according to the present invention without undue experimentation, based on the teaching and guidance presented herein.

Known methods of DNA or RNA amplification include, but are not limited to, polymerase chain reaction (PCR) and related amplification processes (see, e.g., U.S. Patent Nos. 4,683,195, 4,683,202, 4,800,159, 4,965,188, to Mullis, *et al.*; 4,795,699 and 4,921,794 to Tabor, *et al.*; 5,142,033 to Innis; 5,122,464 to Wilson, *et al.*; 5,091,310 to Innis; 5,066,584 to Gyllensten, *et al.*; 4,889,818 to Gelfand, *et al.*; 4,994,370 to Silver, *et al.*; 4,766,067 to Biswas; 4,656,134 to Ringold) and RNA mediated amplification which

uses anti-sense RNA to the target sequence as a template for double-stranded DNA synthesis (U.S. Patent No. 5,130,238 to Malek, et al, with the tradename NASBA), the entire contents of which are herein incorporated by reference.

For instance, polymerase chain reaction (PCR) technology can be used to amplify the sequences of polynucleotides of the present invention and related genes directly from genomic DNA or cDNA libraries. PCR and other in vitro amplification methods may also be useful, for example, to clone nucleic acid sequences that code for proteins to be expressed, to make nucleic acids to use as probes for detecting the presence of the desired mRNA in samples, for nucleic acid sequencing, or for other purposes. Examples of techniques sufficient to direct persons of skill through in vitro amplification methods are found in Berger, Sambrook, and Ausubel, as well as Mullis, *et al.*, U.S. Patent No. 4,683,202 (1987); and Innis, *et al.*, *PCR Protocols A Guide to Methods and Applications*, Eds., Academic Press Inc., San Diego, CA (1990).

Synthetic Methods for Constructing Nucleic Acids

The isolated nucleic acids of the present invention can also be prepared by direct chemical synthesis by methods such as the phosphotriester method of Narang, *et al.*, Meth. Enzymol. 68:90-99 (1979); the phosphodiester method of Brown, *et al.*, Meth. Enzymol. 68:109-151 (1979); the diethylphosphoramidite method of Beaucage, *et al.*, Tetra. Letts. 22:1859-1862 (1981); the solid phase phosphoramidite triester method described by Beaucage and Caruthers, Tetra.

Letts. 22(20):1859-1862 (1981), e.g., using an automated synthesizer, e.g., as described in Needham-VanDevanter, *et al.*, Nucleic Acids Res. 12:6159-6168 (1984); and the solid support method of U.S. Patent No. 4,458,066. Chemical synthesis generally produces a single-stranded oligonucleotide, which may be converted into double-stranded DNA by hybridization with a complementary sequence, or by polymerization with a DNA polymerase using the single strand as a template. One of skill in the art will recognize that while chemical synthesis of DNA is limited to sequences of about 100 bases, longer sequences may be obtained by the ligation of shorter sequences.

Recombinant Expression Cassettes

The present invention further provides recombinant expression cassettes comprising a nucleic acid of the present invention. A nucleic acid sequence of the present invention, for example a cDNA or a genomic sequence encoding a full-length polypeptide of the present invention, can be used to construct a recombinant expression cassette which can be introduced into a desired host cell. A recombinant expression cassette will typically comprise a polynucleotide of the present invention operably linked to transcriptional initiation regulatory sequences that will direct the transcription of the polynucleotide in the intended host cell.

Both heterologous and non-heterologous (i.e., endogenous) promoters can be employed to direct expression of the nucleic acids of the present invention. These promoters can also be used, for example, in recombinant expression

cassettes to drive expression of antisense nucleic acids to reduce, increase, or alter F8H content and/or composition in a desired tissue.

In some embodiments, isolated nucleic acids which serve as promoter or enhancer elements can be introduced in the appropriate position (generally upstream) of a non-heterologous form of a polynucleotide of the present invention so as to modulate expression of a polynucleotide of the present invention. For example, endogenous promoters can be altered in vivo or in vitro by mutation, deletion and/or substitution.

A polynucleotide of the present invention can be expressed in either sense or anti-sense orientation, as desired. It will be appreciated that control of gene expression in either sense or anti-sense orientation can have a direct impact on the observable characteristics.

A variety of cross-linking agents, alkylating agents and radical generating species as pendant groups on polynucleotides of the present invention can be used to bind, label, detect and/or cleave nucleic acids. For example, V. Vlassov, *et al.*, Nucleic Acids Res. 14:4065-4076 (1986), describe covalent bonding of a single-stranded DNA fragment with alkylating derivatives of nucleotides complementary to target sequences. A report of similar work by the same group is that by D. G. Knorre, *et al.*, Biochimie 67:785-789 (1985). Iverson and Dervan also showed sequence-specific cleavage of single-stranded DNA mediated by incorporation of a modified nucleotide (J. Am. Chem. Soc. 109:1241-1243 (1987)). A photoactivated crosslinking to single-stranded

oligonucleotides mediated by psoralen was disclosed by B. L. Lee, *et al.*, Biochemistry 27:3197-3203 (1988). Use of crosslinking in triple-helix forming probes was also disclosed by Home, *et al.*, J. Am. Chem. Soc. 112:2435-2437 (1990). Use of N4, N4-ethanocytosine as an alkylating agent to crosslink to single-stranded oligonucleotides has also been described by Webb and Matteucci, J. Am. Chem. Soc. 108:2764-2765 (1986); Nucleic Acids Res. 14:7661-7674 (1986); Feteritz, *et al.*, J. Am. Chem. Soc. 113:4000 (1991). Various compounds are known in the art to bind, detect, label, and/or cleave nucleic acids. See, for example, U.S. Patent Nos. 5,543,507; 5,672,593; 5,484,908; 5,256,648; and 5,681941.

VECTORS AND HOST CELLS

The present invention also relates to vectors that include isolated nucleic acid molecules of the present invention, host cells that are genetically engineered with said recombinant vectors, and the production of F8H polypeptides or fragments thereof by recombinant techniques, as is well known in the art. See, e.g., Sambrook, *et al.*, 1989; Ausubel, *et al.*, 1987-1989, each entirely incorporated herein by reference.

The polynucleotides can optionally be joined to a vector containing a selectable marker for propagation in a host. Generally, a plasmid vector is introduced in a precipitate, such as a calcium phosphate precipitate, or in a complex with a charged lipid. If the vector is a virus, it can be packaged in vitro using an appropriate packaging cell line and then transduced into host cells.

The DNA insert should be operatively linked to an appropriate promoter, such as the phage lambda PL promoter, the *E. coli lac*, *trp* and *tac* promoters, the SV40 early and late promoters and promoters of retroviral LTRs, to name a few. Other suitable promoters will be known to the skilled artisan. The expression constructs will further contain sites for transcription initiation, termination and, in the transcribed region, a ribosome binding site for translation. The coding portion of the mature transcripts expressed by the constructs will preferably include a translation initiating at the beginning and a termination codon appropriately positioned at the end of the mRNA to be translated.

Expression vectors will preferably include at least one selectable marker. Such markers include, e.g., dihydrofolate reductase or neomycin resistance for eucaryotic cell culture, and tetracycline or ampicillin resistance genes for culturing in E. coli and other bacteria. Representative examples of appropriate hosts include, but are not limited to, bacterial cells, such as E. coli, Streptomyces and Salmonella typhimurium cells; fungal cells, such as yeast; insect cells such as Drosophila S2 and Spodoptera Sf9; animal cells such as CHO, COS and Bowes melanoma; and plant cells. Appropriate culture media and conditions for the above-described host cells are known in the art. Vectors preferred for use in bacteria include pQE70, pQE60 and pQE-9, available from Qiagen; pBS vectors, Phagescript vectors, Bluescript vectors, pNH8A, pNH16a, pNH18A, pNH46A, available from Stratagene; and ptrc99a,

pKK223-3, pKK233-3, pDR540, pRIT5 available from Pharmacia. Preferred eucaryotic vectors include pWLNEO, pSV2CAT, pOG44, pXT1 and pSG available from Stratagene; and pSVK3, pBPV, pMSG and pSVL available from Pharmacia. Other suitable vectors will be readily apparent to the skilled artisan.

The polypeptide(s) of the present invention can be expressed in a modified form, such as a fusion protein, and can include not only secretion signals, but also additional heterologous functional regions. For instance, a region of additional amino acids, particularly charged amino acids, can be added to the N-terminus or other region of a polypeptide to improve stability and persistence in the host cell, during purification, or during subsequent handling and storage. For example, several histidine residues may be incorporated to facilitate rapid style, step purification as described in U.S. Patent 4,569,794, herein incorporated by reference. Also, peptide moieties can be added to a polypeptide to facilitate purification. Such regions can be removed prior to final purification of a polypeptide. Such methods are described in many standard laboratory manuals, such as Sambrook, supra, Chapters 17.29-17.42 and 18.1-18.74; Ausubel, supra, Chapters 16, 17 and 18.

Expression of Proteins in Host Cells

Using nucleic acids of the present invention, one may express a protein of the present invention in a recombinantly engineered cell such as bacteria, yeast, insect, or mammalian cell.

In brief summary, the expression of isolated nucleic acids encoding a protein of the present invention will

typically be achieved by operably linking, for example, the DNA or cDNA to a promoter (which is either constitutive or inducible), followed by incorporation into an expression vector. The vectors can be suitable for replication and/or integration in either prokaryotes or eukaryotes. Typical expression vectors contain transcription and translation terminators, initiation sequences and promoters useful for regulation of the expression of the DNA encoding a protein of the present invention. To obtain high level expression of a cloned gene, it is desirable to construct expression vectors which contain a strong promoter to direct transcription, a ribosome binding site for translational initiation, and a transcription/translation terminator.

One of skill would recognize that modifications can be made to a protein of the present invention without diminishing its biological activity. Some modifications may be made to facilitate the cloning, expression, or incorporation of the targeting molecule into a fusion protein. Such modifications are well known to those of skill in the art and include, for example, a methionine added at the amino terminus to provide an initiation site, or additional amino acids (e.g., poly His) placed on either terminus to create conveniently located restriction sites or termination codons or purification sequences.

Expression in Prokaryotes

Prokaryotic cells may be used as hosts for expression, for example, various strains of E. coli, or other microbial strains. Commonly used prokaryotic control sequences which are defined herein to include promoters for transcription

initiation, optionally with an operator, along with ribosome binding sites, include such commonly used promoters as the beta lactamase (penicillinase) and lactose (lac) promoter systems (Chang, *et al.*, Nature 198:1056 (1977)), the tryptophan (trp) promoter system (Goeddel, *et al.*, Nucleic Acids Res. 8:4057 (1980)), and the lambda derived P L promoter and N-gene ribosome binding site (Shimatake, *et al.*, Nature 292:128 (1981)). The inclusion of selection markers in DNA vectors transfected in E. coli is also useful. Examples of such markers include genes specifying resistance to ampicillin, tetracycline, or chloramphenicol.

The vector is selected to allow introduction into the appropriate host cell. Bacterial vectors are typically of plasmid or phage origin. Appropriate bacterial cells are infected with phage vector particles or transfected with naked phage vector DNA. If a plasmid vector is used, the bacterial cells are transfected with the plasmid vector DNA. Expression systems for expressing a protein of the present invention are available using Bacillus sp. and Salmonella (Palva, *et al.*, Gene 22:229-235 (1983); Mosbach, *et al.*, Nature 302:543-545 (1983)).

Expression in Eukaryotes

A variety of eukaryotic expression systems such as yeast, insect cell lines, plant and mammalian cells, are known to those of skill in the art. As explained briefly below, a nucleic acid of the present invention can be expressed in these eukaryotic systems.

Synthesis of heterologous proteins in yeast is well known. See~~, e.g.,~~ F. Sherman, ~~*et al., Methods*~~ *in Yeast*

Genetics, Cold Spring Harbor Laboratory (1982). Two widely utilized yeast for production of eukaryotic proteins are Saccharomyces cerevisiae and Pichia pastoris. Vectors, strains, and protocols for expression in Saccharomyces and Pichia are known in the art and available from commercial suppliers (e.g., Invitrogen). Suitable vectors usually have expression control sequences, such as promoters, including 3-phosphoglycerate kinase or alcohol oxidase, and an origin of replication, termination sequences and the like as desired.

The sequences encoding proteins of the present invention can also be ligated to various expression vectors for use in transfecting cell cultures of, for instance, mammalian, insect, or plant origin. Mammalian cell systems often will comprise cell monolayers, although cell suspensions may also be used. A number of suitable host cell lines have been developed in the art, and include the HEK293, BHK21, and CHO cell lines. Expression vectors for these cells can include expression control sequences, such as an origin of replication, a promoter (e.g., the CMV promoter, a HSV tk promoter or pgk (phosphoglycerate kinase) promoter), an enhancer (Queen, *et al.*, Immunol. Rev. 89:49 (1986)), and sites such as ribosome binding sites, RNA splice sites, polyadenylation sites (e.g., an SV40 large T Ag poly A addition site), and transcriptional terminator sequences. Other animal cells useful for production of proteins of the present invention are available, for instance, from the American Type Culture Collection Catalogue of Cell Lines and Hybridomas (7th edition, 1992).

Appropriate vectors for expressing proteins of the present invention in insect cells are usually derived from the SF9 baculovirus. Suitable insect cell lines include mosquito larvae, silkworm, armyworm, moth and Drosophila cell lines such as a Schneider cell line (See Schneider, J. Embryol. Exp. Morphol. 27:353-365 (1987).

As with yeast, when higher animal or plant host cells are employed, polyadenlyation or transcription terminator sequences are typically incorporated into the vector. An example of a terminator sequence is the polyadenlyation sequence from the bovine growth hormone gene. Sequences for accurate splicing of the transcript may also be included. An example of a splicing sequence is the VP1 intron from SV40 (Sprague, *et al.*, J. Virol. 45:773-781 (1983)). Additionally, gene sequences to control replication in the host cell may be incorporated into the vector such as those found in bovine papilloma virus type-vectors. M. Saveria-Campo, *Bovine Papilloma Virus DNA, a Eukaryotic Cloning Vector in DNA Cloning Vol. II, a Practical Approach*, D. M. Glover, Ed., IRL Press, Arlington, VA, pp. 213-238 (1985).

Protein Purification

A F8H polypeptide can be recovered and purified from recombinant cell cultures by well-known methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Most preferably, high performance liquid chromatography ("HPLC") is employed for purification.

Polypeptides of the present invention include naturally purified products, products of chemical synthetic procedures, and products produced by recombinant techniques from a prokaryotic or eucaryotic host, including, for example, bacterial, yeast, higher plant, insect and mammalian cells. Depending upon the host employed in a recombinant production procedure, the polypeptides of the present invention can be glycosylated ~~or~~ can be non-glycosylated. In addition, polypeptides of the invention can also include an initial modified methionine residue, in some cases as a result of host-mediated processes. Such methods are described in many standard laboratory manuals, such as Sambrook, supra, Chapters 17.37-17.42; Ausubel, supra, Chapters 10, 12, 13, 16, 18 and 20.

***F8H* POLYPEPTIDES, FRAGMENTS, AND VARIANTS**

The invention further provides an isolated F8H polypeptide comprising a fragment or specified variant of the amino acid sequence encoded by the deposited cDNAs, or the amino acid sequence in SEQ ID NO:2.

The isolated proteins of the present invention comprise a polypeptide having at least 5 - 10 amino acids encoded by any one of the polynucleotides of the present invention as discussed more fully, supra, or polypeptides which are conservatively modified variants thereof.

Exemplary polypeptide sequences are provided by SEQ ID NO:2. The proteins of the present invention or variants thereof can comprise any number of contiguous amino acid residues from a polypeptide of the present invention, wherein the number of residues is selected from the group of integers

consisting of from 10 to the number of residues in a full-length F8H polypeptide. Optionally, this subsequence of contiguous amino acids is at least 15, 20, 25, 30, 35, or 40 amino acids in length, often at least 50, 60, 70, 80, or 90 amino acids in length. Further, the number of such subsequences can be any integer selected from the group consisting of from 1 to 20, such as 2, 3, 4, or 5.

As those of skill will appreciate, the present invention includes biologically active polypeptides of the present invention (i.e., enzymes). Biologically active polypeptides have a specific activity at least 20%, 30%, or 40%, and preferably at least 50%, 60%, or 70%, and most preferably at least 80%, 90%, or 95% - 100% of that of the native (non-synthetic), endogenous polypeptide. Further, the substrate specificity (k_{cat}/K_m) is optionally substantially similar to the native (non-synthetic), endogenous polypeptide. Typically, the K_m will be at least 30%, 40%, or 50%, that of the native (non-synthetic), endogenous polypeptide; and more preferably at least 60%, 70%, 80%, or 90%. Methods of assaying and quantifying measures of enzymatic activity and substrate specificity (k_{cat}/K_m), are well known to those of skill in the art.

Generally, the polypeptides of the present invention will, when presented as an immunogen, elicit production of an antibody specifically reactive to a polypeptide of the present invention encoded by a polynucleotide of the present invention as described, supra. Exemplary polypeptides include those which are full-length, such as those disclosed herein. Further, the proteins of the present invention will

not bind to antisera raised against a polypeptide of the present invention which has been fully immunosorbed with the same polypeptide. Immunoassays for determining binding are well known to those of skill in the art. A preferred immunoassay is a competitive immunoassay as discussed, infra. Thus, the proteins of the present invention can be employed as immunogens for constructing antibodies immunoreactive to a protein of the present invention for such exemplary utilities as immunoassays or protein purification techniques.

A F8H polypeptide of the present invention can include one or more amino acid substitutions, deletions or additions, either from natural mutations or human manipulation, as specified herein.

Of course, the number of amino acid substitutions a skilled artisan would make depends on many factors, including those described above. Generally speaking, the number of amino acid substitutions for any given F8H polypeptide will not be more than 20, 10, 5, or 3, such as 1-20 or any range or value therein.

Amino acids in an F8H polypeptide of the present invention that are essential for function can be identified by methods known in the art, such as site-directed mutagenesis, or alanine-scanning mutagenesis (Cunningham and Wells, Science 244:1081-1085 (1989)). The latter procedure introduces single alanine mutations at every residue in the molecule. The resulting mutant molecules are then tested for biological activity. Sites that are critical for ligand-protein binding can also be identified by structural analysis such as crystallization, nuclear magnetic resonance

or photoaffinity labeling (Smith, *et al.*, J. Mol. Biol. 224:899-904 (1992) and de Vos, *et al.*, Science 255:306-312 (1992)).

F8H polypeptides of the present invention can include, but are not limited to, at least one selected from extracellular, intracellular, trans-membrane and active versions of SEQ ID NO:2.

Antigenic/Epitope Comprising F8H Peptide and Polypeptides

In another aspect, the invention provides a peptide or polypeptide comprising an epitope-bearing portion of a polypeptide of the invention according to methods well known in the art. See, e.g., Colligan, et al,. ed., Current *Protocols in Immunology*, Greene Publishing, NY (1993-1998), Ausubel, supra, entirely incorporated herein by reference.

An "immunogenic epitope" can be defined as a part of a polypeptide that elicits an antibody response when the whole polypeptide is the immunogen. On the other hand, a region of a polypeptide molecule to which an antibody can bind is defined as an "antigenic epitope." The number of immunogenic epitopes of a polypeptide generally is less than the number of antigenic epitopes. See, for instance, Geysen, *et al.*, Proc. Natl. Acad. Sci. USA 81:3998-4002 (1983).

As to the selection of peptides or polypeptides bearing an antigenic epitope, it is well known in the art that relatively short synthetic peptides that mimic part of a polypeptide sequence are routinely capable of eliciting an antiserum that reacts with the partially mimicked polypeptide. See, for instance, J. G. Sutcliffe, *et al.*,

"Antibodies that react with preidentified sites on polypeptides," Science 219:660-666 (1983).

Antigenic epitope-bearing peptides and polypeptides of the invention are useful to raise antibodies, including monoclonal antibodies, or screen antibodies, including fragments or single chain antibodies, that bind specifically to a polypeptide of the invention. See, for instance, Wilson, *et al.*, Cell 37:767-778 (1984) at 777. Antigenic epitope-bearing peptides and polypeptides of the invention preferably contain a sequence of at least seven, more preferably at least nine, and most preferably between at least about 15 to about 30 amino acids contained within the amino acid sequence of a polypeptide of the invention.

The epitope-bearing peptides and polypeptides of the invention can be produced by any conventional means. R. A. Houghten, "General method for the rapid solid-phase synthesis of large numbers of peptides: specificity of antigen-antibody interaction at the level of individual amino acids," Proc. Natl. Acad. Sci. USA 82:5131-5135 (1985). This "Simultaneous Multiple Peptide Synthesis (SMPS)" process is further described in U.S. Patent No. 4,631,211 to Houghten, *et al.* (1986).

As one of skill in the art will appreciate, F8H polypeptides of the present invention and the epitope-bearing fragments thereof can be combined with parts of the constant domain of immunoglobulins (IgG), resulting in chimeric polypeptides. These fusion proteins facilitate purification and show an increased half-life in vivo. Fusion proteins that have a disulfide-linked dimeric

structure due to the IgG part can also be more efficient in binding and neutralizing other molecules than the monomeric F8H polypeptide or polypeptide fragment alone (Fountoulakis, *et al.*, J. Biochem. 270:3958-3964 (1995)).

Production of Antibodies

The polypeptides of this invention and fragments thereof may be used in the production of antibodies. The term "antibody" as used herein describes antibodies, fragments of antibodies (such as, but not limited, to Fab, Fab', Fab2', and Fv fragments), and modified versions thereof, as well known in the art (e.g., chimeric, humanized, recombinant, veneered, resurfaced or CDR-grafted) such antibodies are capable of binding antigens of a similar nature as the parent antibody molecule from which they are derived. The instant invention also encompasses single chain polypeptide binding molecules.

The production of antibodies, both monoclonal and polyclonal, in animals is well known in the art. See, e.g., Colligan, supra, entirely incorporated herein by reference.

Single chain antibodies and libraries thereof are yet another variety of genetically engineered antibody technology that is well known in the art. (See, e.g., R. E. Bird, *et al.*, Science 242:423-426 (1988); PCT Publication Nos. WO 88/01649, WO 90/14430, and WO 91/10737. Single chain antibody technology involves covalently joining the binding regions of heavy and light chains to generate a single polypeptide chain. The binding specificity of the intact antibody molecule is thereby reproduced on a single polypeptide chain.

Antibodies included in this invention are useful in diagnostics, therapeutics or in diagnostic/therapeutic combinations.

The polypeptides of this invention or suitable fragments thereof can be used to generate polyclonal or monoclonal antibodies, and various inter-species hybrids, or humanized antibodies, or antibody fragments, or single-chain antibodies. The techniques for producing antibodies are well known to skilled artisans. (See, e.g., Colligan supra; *Monoclonal Antibodies: Principles & Applications*, Ed. J. R. Birch & E. S. Lennox, Wiley-Liss (1995).

A polypeptide used as an immunogen may be modified or administered in an adjuvant, by subcutaneous or intraperitoneal injection into, for example, a mouse or a rabbit. For the production of monoclonal antibodies, spleen cells from immunized animals are removed, fused with myeloma or other suitable known cells, and allowed to become monoclonal antibody producing hybridoma cells in the manner known to the skilled artisan. Hybridomas that secrete a desired antibody molecule can be screened by a variety of well known methods, for example ELISA assay, Western blot analysis, or radioimmunoassay (Lutz, *et al.* Exp. Cell Res. 175:109-124 (1988); *Monoclonal Antibodies: Principles & Applications*, Ed. J. R. Birch & E. S. Lennox, Wiley-Liss (1995); Colligan, supra).

For some applications labeled antibodies are desirable. Procedures for labeling antibody molecules are widely known, including for example, the use of radioisotopes, affinity labels, such as biotin or avidin, enzymatic labels, for

example horseradish peroxidase, and fluorescent labels, such as FITC or rhodamine (See, e.g., Colligan, supra).

Labeled antibodies are useful for a variety of diagnostic applications. In one embodiment the present invention relates to the use of labeled antibodies to detect the presence of an F8H polypeptide. Alternatively, the antibodies could be used in a screen to identify potential modulators of an F8H polypeptide. For example, in a competitive displacement assay, the antibody or compound to be tested is labeled by any suitable method. Competitive displacement of an antibody from an antibody-antigen complex by a test compound such that a test compound-antigen complex is formed provides a method for identifying compounds that bind HPLFP.

Transgenics and Chimeric Non-Human Mammals

The present invention is also directed to a transgenic non-human eukaryotic animal (preferably a rodent, such as a mouse) the germ cells and somatic cells of which contain genomic or cDNA according to the present invention which codes for an F8H polypeptide. At least one F8H nucleic acid can be introduced into the animal to be made transgenic, or an ancestor of the animal, at an embryonic stage, preferably the 1-1000 cell or oocyte, stage, and preferably not later than about the 64-cell stage. The term "transgene," as used herein, means a gene which is incorporated into the genome of the animal and is expressed in the animal, resulting in the presence of at least one F8H polypeptide in the transgenic animal.

There are several means by which such an F8H nucleic acid can be introduced into a cell or genome of the animal embryo so as to be chromosomally incorporated and expressed according to known methods.

Chimeric non-human mammals in which fewer than all of the somatic and germ cells contain the F8H polypeptide nucleic acid of the present invention are also intended to be within the scope of the present invention.

Chimeric, non-human, mammals having human cells or tissues engrafted therein are also encompassed by the present invention. Methods for producing chimeric non-human mammals are provided, e.g., in U.S. Serial Nos. 07/508,225, 07/518,748, 07/529,217, 07/562,746, 07/596,518, 07/574,748, 07/575,962, 07/207,273, 07/241,590 and 07/137,173, which are entirely incorporated herein by reference.

The techniques described in Leder, U.S. Patent No. 4,736,866 (hereby entirely incorporated by reference) for producing transgenic non-human mammals may be used in the present invention. The various techniques described in U.S. Patent Nos. 5,545,807, 5,073,490, 5,347,075, and 4,736,866, the entire contents of which are hereby incorporated by reference, may also be used.

Transgenic animals expressing an F8H polypeptide and/or nucleic acid can be used to test compounds and treatment modalities to suppress or cure a pathology relating to the F8H polypeptide or F8H nucleic acid. Such transgenic animals will also serve as a model for testing diagnostic methods for the same diseases. Transgenic animals according to the

present invention can also be used as a source of cells for cell culture.

Having generally described the invention, the same will be more readily understood by reference to the following examples, which are provided by way of illustration and are not intended as limiting.

Example 1: Expression and Purification of an F8H Polypeptide in E. coli

The bacterial expression vector pQE60 is used for bacterial expression in this example. (QIAGEN, Inc., Chatsworth, CA). pQE60 encodes ampicillin antibiotic resistance ("Amp®") and contains a bacterial origin of replication ("ori"), an IPTG inducible promoter, a ribosome binding site ("RBS"), six codons encoding histidine residues that allow affinity purification using nickel-nitrilo-tri-acetic acid ("Ni-NTA") affinity resin (QIAGEN, Inc.), and suitable single restriction enzyme cleavage sites. These elements are arranged such that a DNA fragment encoding an F8H polypeptide can be inserted in such a way as to produce an F8H polypeptide with six His residues (i.e., a "6 X His tag") covalently linked to the amino or carboxyl terminus of the polypeptide. However, a polypeptide coding sequence can optionally be inserted such that translation of the six His codons is prevented and, therefore, a polypeptide is produced with no 6 X His tag.

The nucleic acid sequence encoding the desired portion of an F8H polypeptide lacking the hydrophobic leader

sequence is amplified from the deposited cDNA or from mRNA, or cDNA from a suitable tissue clone using PCR oligonucleotide primers (based on the sequences presented, e.g., as presented in SEQ ID NO:1, which anneal to the 5' end of the desired portion of an F8H nucleic acid and to sequences at the 3' end. Additional nucleotides containing restriction sites to facilitate cloning in the pQE60 vector are added to the 5' and 3' sequences, respectively.

For cloning an F8H polypeptide, the 5' and 3' primers have nucleotides corresponding or complementary to a portion of the coding sequence of an F8H nucleic acid, e.g., as presented in SEQ ID NO:1, according to known methods. One of ordinary skill in the art would appreciate that the point in a polypeptide coding sequence where the 5' primer begins can be varied to amplify a desired portion of the complete polypeptide, shorter or longer than the mature form.

The amplified F8H nucleic acid fragments and the vector pQE60 are digested with appropriate restriction enzymes and the digested DNAs are then ligated together. Insertion of the F8H DNA into the restricted pQE60 vector places an F8H polypeptide coding region including its associated stop codon downstream from an IPTG-inducible promoter and in-frame with an initiating ATG codon. The associated stop codon prevents translation of the six histidine codons downstream of the insertion point.

The ligation mixture is transformed into competent E. coli cells using standard procedures such as those described in Sambrook, *et al.*, 1989; Ausubel, 1987-1998. E. coli strain M15/rep4, containing multiple copies of the plasmid

pREP4, which expresses the lac repressor and confers kanamycin resistance, is used in carrying out the illustrative example described herein. This strain, which is only one of many that are suitable for expressing F8H polypeptide, is available commercially from QIAGEN, Inc. Transformants are identified by their ability to grow on LB plates in the presence of ampicillin and kanamycin. Plasmid DNA is isolated from resistant colonies and the identity of the cloned DNA confirmed by restriction analysis, PCR and DNA sequencing.

Clones containing the desired constructs are grown overnight ("O/N") in LB medium supplemented with ampicillin (100 μg/ml) and kanamycin (25 μg/ml). Isopropyl-b-D-thiogalactopyranoside ("IPTG") is added to a fresh culture to induce transcription from the lac repressor sensitive promoter. Cells subsequently are incubated further for 3 to 4 hours and harvested by centrifugation.

The cells are lysed by any suitable method, and the supernatant containing the F8H is dialyzed against 50 mM Na-acetate buffer pH6, supplemented with 200 mM NaCl. The polypeptide is purified by ion exchange, hydrophobic interaction and size exclusion chromatography. Alternatively, an affinity chromatography step such as an antibody column is used to obtain pure F8H polypeptide. The purified polypeptide is stored at 4°C or frozen at -40°C to -120°C.

Example 2: Cloning and Expression of an F8H Polypeptide in a Baculovirus Expression System

In this illustrative example, the plasmid shuttle vector pA2 GP is used to insert the cloned DNA encoding the mature polypeptide into a baculovirus to express an F8H polypeptide, using standard methods as described in Summers, *et al.*, A Manual of Methods for Baculovirus Vectors and Insect Cell Culture Procedures, Texas Agricultural Experimental Station Bulletin No. 1555 (1987). This expression vector contains the Autographa californica nuclear polyhedrosis virus promoter (AcMNPV) followed by the secretory signal peptide (leader) of the baculovirus gp67 polypeptide and convenient restriction sites. The polyadenylation site of the simian virus 40 ("SV40") is used for efficient polyadenylation. For easy selection of recombinant virus, the plasmid contains the beta-galactosidase gene from E. coli under control of a weak Drosophila promoter in the same orientation, followed by the polyadenylation signal of the polyhedrin gene. Inserted genes are flanked on both sides by viral sequences for cell-mediated homologous recombination with wild-type viral DNA to generate viable virus that express the cloned polynucleotide.

Other baculovirus vectors may be used in place of the vector above, such as pAc373, pVL941 and pAcIM1, as long as the construct provides appropriately located signals for transcription, translation, secretion and the like, including a signal peptide and an in-frame ATG, as required.

Such vectors are described, for instance, in Luckow, *et al.*, Virology 170:31-39.

A cDNA sequence encoding the mature F8H polypeptide, lacking the ATG initiation codon and naturally associated nucleotide binding site(s), is amplified using PCR oligonucleotide primers corresponding to the 5' and 3' sequences of the gene. Non-limiting examples include 5' and 3' primers having nucleotides corresponding or complementary to a portion of the coding sequence of SEQ ID NO:1, according to known methods.

The amplified fragment is isolated and digested with an appropriate restriction enzyme, and designated herein "F1".

The plasmid vector is digested with the corresponding restriction enzymes and optionally, can be dephosphorylated using routine procedures. This vector is designated herein "V1".

Fragment F1 and plasmid V1 are ligated together, and E. coli HB101 or other suitable E. coli hosts such as XL-1 Blue (Stratagene Cloning Systems, La Jolla, CA) are transformed with the ligation mixture and spread on culture plates. Bacteria that contain a plasmid with the human F8H gene can be identified using the PCR method. This plasmid is designated herein pBac F8H.

Five µg of the plasmid pBacF8H is co-transfected with 1.0 µg of a commercially available linearized baculovirus DNA ("BaculoGold™ baculovirus DNA", Pharmingen, San Diego, CA), using the lipofection method described by Felgner, *et al.*, Proc. Natl. Acad. Sci. USA 84:7413-7417 (1987). 1 µg of BaculoGold™ virus DNA and 5 µg of the plasmid pBac F8H

are mixed in a sterile well of a microtiter plate containing 50 µl of serum-free Grace's medium (Life Technologies, Inc., Rockville, MD). Afterwards, 10 µl Lipofectin plus 90 µl Grace's medium are added, mixed and incubated for 15 minutes at room temperature. Then the transfection mixture is added drop-wise to Sf9 insect cells (ATCC CRL 1711) seeded in a 35 mm tissue culture plate with 1 ml Grace's medium without serum. The plate is rocked back and forth to mix the newly added solution. The plate is then incubated for 5 hours at 27°C. After 5 hours the transfection solution is removed from the plate and 1 ml of Grace's medium supplemented with 10% fetal calf serum is added. The plate is put back into an incubator and cultivation is continued at 27°C for four days.

After four days the supernatant is collected and a plaque assay is performed, according to known methods. An agarose gel with "Blue Gal" (Life Technologies, Inc., Rockville, MD) is used to allow easy identification and isolation of gal-expressing clones, which produce blue-stained plaques. (A detailed description of a "plaque assay" of this type can also be found in the user's guide for insect cell culture and baculovirology distributed by Life Technologies, Inc., Rockville, MD, page 9-10). After appropriate incubation, blue stained plaques are picked with a micropipettor tip. Recombinant viruses are resuspended in a microcentrifuge tube containing 200 µl of Grace's medium and the suspension is used to infect Sf9 cells seeded in 35 mm dishes. Four days later the supernatants of these

culture dishes are harvested and stored at 4°C. The recombinant virus is called V-F8H.

To verify the expression of the F8H gene, Sf9 cells are grown in Grace's medium supplemented with 10% heat - inactivated FBS. The cells are infected with the recombinant baculovirus V-F8H at a multiplicity of infection ("MOI") of about 2. Six hours later the medium is removed and replaced with SF900 II medium minus methionine and cysteine (available from Life Technologies, Inc., Rockville, MD). If radiolabeled polypeptides are desired, 42 hours later, 5 mCi of ^{35}S-methionine and 5 mCi ^{35}S-cysteine (available from Amersham) are added. The cells are further incubated for 16 hours and then harvested by centrifugation. The polypeptides in the supernatant as well as intracellular polypeptides are analyzed by SDS-PAGE followed by autoradiography (if radiolabeled). Microsequencing of the amino acid sequence of the amino terminus of purified polypeptide can be used to determine the amino terminal sequence of the mature polypeptide and thus the cleavage point and length of the secretory signal peptide.

Example 3: Cloning and Expression of F8H in Mammalian Cells

Suitable expression vectors for use in practicing the present invention include, for example, vectors such as pIRES1neo, pRetro-Off, pRetro-On, PLXSN, or pLNCX (Clonetech Labs, Palo Alto, CA), pcDNA3.1 (+/-), pcDNA/Zeo (+/-) or pcDNA3.1/Hygro (+/-) (Invitrogen), PSVL and PMSG (Pharmacia, Uppsala, Sweden), pRSVcat (ATCC 37152), pSV2dhfr (ATCC 37146) and pBC12MI (ATCC 67109). Mammalian host cells that

could be used include human Hela 293, H9 and Jurkat cells, mouse NIH3T3 and C127 cells, Cos 1, Cos 7 and CV 1, quail QC1-3 cells, mouse L cells and Chinese hamster ovary (CHO) cells.

Alternatively, the F8H gene can be expressed in a stable cell line that contains the gene integrated into a chromosome by co-transfection with a selectable marker such as dhfr, gpt, neomycin, or hygromycin allows.

The expression vectors pC1 and pC4 contain the strong promoter (LTR) of the Rous Sarcoma Virus (Cullen, *et al.*, Molec. Cell. Biol. 5:438-447 (1985)) plus a fragment of the CMV-enhancer (Boshart, *et al.*, Cell 41:521-530 (1985)). Multiple cloning sites facilitate the cloning of the gene of interest. The vectors contain in addition the 3' intron, the polyadenylation and termination signal of the rat preproinsulin gene.

Example 3(a): Cloning and Expression in COS Cells

The expression plasmid, pF8H HA, is made by cloning a cDNA encoding F8H into the expression vector pcDNAI/Amp or pcDNAIII (available from Invitrogen, Inc.).

The expression vector pcDNAI/amp contains: (1) an E. coli origin of replication effective for propagation in E. coli and other prokaryotic cells; (2) an ampicillin resistance gene for selection of plasmid-containing prokaryotic cells; (3) an SV40 origin of replication for propagation in eucaryotic cells; (4) a CMV promoter, a polylinker, an SV40 intron; (5) several codons encoding a hemagglutinin fragment (i.e., an "HA" tag to facilitate

purification) or HIS tag (see, e.g, Ausubel, supra) followed by a termination codon and polyadenylation signal arranged so that a cDNA can be conveniently placed under expression control of the CMV promoter and operably linked to the SV40 intron and the polyadenylation signal by means of restriction sites in the polylinker.

A DNA fragment encoding the F8H is cloned into the polylinker region of the vector so that recombinant polypeptide expression is directed by the CMV promoter. The F8H cDNA of the deposited clone or other source is amplified using primers that contain convenient restriction sites. Non-limiting examples of suitable primers include those based on the coding sequences presented in SEQ ID NO:1.

The PCR amplified DNA fragment and the vector, pcDNAI/Amp, are digested with suitable restriction enzyme(s) and then ligated. The ligation mixture is transformed into E. coli strain SURE (available from Stratagene Cloning Systems, 11099 North Torrey Pines Road, La Jolla, CA 92037), and the transformed culture is plated on ampicillin plates. Plasmid DNA is isolated from resistant colonies and examined by restriction analysis or other means for the presence of the F8H-encoding fragment.

For expression of recombinant F8H, COS cells are transfected with an expression vector, as described above, using DEAE-DEXTRAN, as described, for instance, in Sambrook, *et al.*, *Molecular Cloning: a Laboratory Manual*, Cold Spring Laboratory Press, Cold Spring Harbor, New York (1989). Cells are incubated under conditions for expression of F8H by the vector.

Expression of the F8H-HA fusion polypeptide is detected by radiolabeling and immunoprecipitation, using methods described in, for example Harlow, *et al.*, *Antibodies: A Laboratory Manual, 2nd Ed.*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York (1988). To this end, two days after transfection, the cells are labeled by incubation in media containing ^{35}S-cysteine for 8 hours. The cells and the media are collected, and the cells are washed and lysed with detergent-containing RIPA buffer: 150 mM NaCl, 1% NP-40, 0.1% SDS, 0.5% DOC, 50 mM TRIS, pH 7.5, as described by Wilson, *et al.* cited above. Proteins are precipitated from the cell lysate and from the culture media using an HA-specific monoclonal antibody. The precipitated polypeptides then are analyzed by SDS-PAGE and autoradiography. An expression product of the expected size is seen in the cell lysate, which is not seen in negative controls.

The vector pC4 is used for the expression of F8H polypeptide. Plasmid pC4 is a derivative of the plasmid pSV2-dhfr (ATCC Accession No. 37146). The plasmid contains the mouse DHFR gene under control of the SV40 early promoter. Chinese hamster ovary- or other cells lacking dihydrofolate activity that are transfected with these plasmids can be selected by growing the cells in a selective medium (alpha minus MEM, Life Technologies) supplemented with methotrexate. The amplification of the DHFR genes in cells resistant to methotrexate (MTX) has been well documented (see, e.g., F. W. Alt, *et al.*, J. Biol. Chem. 253:1357-1370 (1978); J. L. Hamlin and C. Ma, Biochem. et

Biophys. Acta 1097:107-143 (1990); and M. J. Page and M. A. Sydenham, Biotechnology 9:64-68 (1991)). Cells grown in increasing concentrations of MTX develop resistance to the drug by overproducing the target enzyme, DHFR, as a result of amplification of the DHFR gene. If a second gene is linked to the DHFR gene, it is usually co-amplified and over-expressed. It is known in the art that this approach can be used to develop cell lines carrying more than 1,000 copies of the amplified gene(s).

Plasmid pC4 contains the gene of interest, the strong promoter of the long terminal repeat (LTR) of the Rous Sarcoma Virus (Cullen, *et al.*, Molec. Cell. Biol. 5:438-447 (1985)), plus a fragment isolated from the enhancer of the immediate early gene of human cytomegalovirus (CMV) (Boshart, *et al.*, Cell 41:521-530 (1985)). Downstream of the promoter are BamHI, XbaI, and Asp718 restriction enzyme cleavage sites that allow integration of the genes. Behind these cloning sites the plasmid contains the 3' intron and polyadenylation site of the rat preproinsulin gene. Other high efficiency promoters can also be used for the expression, e.g., the human b-actin promoter, the SV40 early or late promoters, or the long terminal repeats from other retroviruses, e.g., HIV and HTLVI. Clontech's Tet-Off and Tet-On gene expression systems and similar systems can be used to express the F8H in a regulated way in mammalian cells (M. Gossen, and H. Bujard, Proc. Natl. Acad. Sci. USA 89: 5547-5551 (1992)). For polyadenylation of the mRNA, other signals, e.g., from the human growth hormone or globin genes can be used as well. Stable cell lines carrying a

gene of interest integrated into a chromosome can also be selected upon co-transfection with a selectable marker such as gpt, G418 or hygromycin. It is advantageous to use more than one selectable marker in the beginning, e.g., G418 plus methotrexate.

The plasmid pC4 is digested with restriction enzymes and then dephosphorylated by procedures known in the art. The vector is then isolated from a 1% agarose gel.

The DNA sequence encoding the complete F8H polypeptide is amplified using PCR oligonucleotide primers corresponding to the 5' and 3' sequences of the gene. Non-limiting examples include 5' and 3′ primers having nucleotides corresponding or complementary to a portion of the coding sequence of an F8H, e.g., as presented in SEQ ID NO:1, according to known methods.

The amplified fragment is digested with suitable endonucleases and then purified again on a 1% agarose gel. The isolated fragment and the dephosphorylated vector are then ligated with T4 DNA ligase. E. coli HB101 or XL-1 Blue cells are then transformed, and bacteria are identified that contain the fragment inserted into plasmid pC4 using, for instance, restriction enzyme analysis.

Chinese hamster ovary (CHO) cells lacking an active DHFR gene are used for transfection. 5 µg of the expression plasmid pC4 is cotransfected with 0.5 µg of the plasmid pSV2-neo using lipofectin. The plasmid pSV2neo contains a dominant selectable marker, the neo gene from Tn5, encoding an enzyme that confers resistance to a group of antibiotics including G418. The cells are seeded in alpha minus MEM

supplemented with 1 μg/ml G418. After 2 days, the cells are trypsinized and seeded in hybridoma cloning plates (Greiner, Germany) in alpha minus MEM supplemented with 10, 25, or 50 ng/ml of methotrexate plus 1 μg/ml G418. After about 10-14 days single clones are trypsinized and seeded in 6-well petri dishes or 10 ml flasks using different concentrations of methotrexate (50 nM, 100 nM, 200 nM, 400 nM, 800 nM). Clones growing at the highest concentrations of methotrexate are transferred to new 6-well plates containing higher concentrations of methotrexate (1 mM, 2 mM, 5 mM, 10 mM, 20 mM). The same procedure is repeated until clones are obtained which grow at a concentration of 100 - 200 mM. Expression of the desired gene product is analyzed, for instance, by SDS-PAGE and Western blot or by reverse phase HPLC analysis.

Example 4: Tissue Distribution of F8H mRNA Expression

Northern blot analysis is carried out to examine F8H gene expression in human tissues, using methods described by, among others, Sambrook, *et al.*, cited above. A cDNA probe containing the entire nucleotide sequence of an F8H polypeptide (SEQ ID NO:1) is labeled with ^{32}P using the Rediprime™ DNA labeling system (Amersham Life Science), according to the manufacturer's instructions. After labeling, the probe is purified using a CHROMA SPIN-100™ column (Clontech Laboratories, Inc.), according to the manufacturer's protocol number PT1200-1. The purified and labeled probe is used to examine various human tissues for F8H mRNA.

Multiple Tissue Northern (MTN) blots containing various human tissues (H) or human immune system tissues (IM) are obtained from Clontech and are examined with the labeled probe using ExpressHyb hybridization solution (Clontech) according to manufacturer's protocol number PT1190-1. Following hybridization and washing, the blots are mounted and exposed to film at -70°C overnight, and films developed according to standard procedures. The results show F8H polypeptides to be selectively expressed in hematopoietic, heart and reproductive and other tissues.

It will be clear that the invention can be practiced otherwise than as particularly described in the foregoing description and examples.

Numerous modifications and variations of the present invention are possible in light of the above teachings and, therefore, are within the scope of the appended claims.

WE CLAIM:

1. An isolated nucleic acid comprising a sequence corresponding or complementary to at least 10 nucleotides of SEQ ID NO:1.

2. An isolated nucleic acid comprising an F8H polynucleotide encoding at least a 20 amino acid fragment of a protein sequence designated SEQ ID NO:2.

3. An isolated polypeptide comprising at least 20 amino acids of the amino acid sequence of SEQ ID NO:2.

4. An isolated polypeptide comprising an isolated polypeptide encoded by a nucleic acid according to any of claims 1-2.

5 A composition comprising an isolated polypeptide according to any of claims 3-4 and a carrier or diluent.

6. A vector comprising a nucleic acid according to any of claims 1-2.

7. A host cell comprising an isolated nucleic acid according to any of claims 1-2.

8. An antibody or at least one fragment thereof that binds an epitope specific to at least 5 contiguous amino acids of at least one F8H polypeptide according to any of claims 3-4.

9. A host cell, expressing at least one antibody or at least one fragment thereof according to claim 8.

10. A method for producing at least one antibody, comprising culturing a host cell according to claim 9.

11. A transgenic or chimeric non-human animal, comprising at least one host cell according to claim 7.

12. A method for identifying compounds that bind at least one F8H polypeptide, comprising

a) admixing at least one isolated F8H polypeptide according to any of claims 3-4 with at least one test compound or composition; and

b) detecting at least one binding interaction between said at least one F8H polypeptide and the test compound or composition.

13. A compound or composition detected by method according to claim 12.

14. Any invention described herein.

SEQUENCE LISTING

<110> Eli Lilly and Company

<120> FACTOR 8 HOMOLOGUE NUCLEIC ACIDS, POLYPEPTIDES,
METHODS, USES

<130> X-11970 PCT

<140>
<141>

<160> 2

<170> MS-DOS Text

<210> 1
<211> 2046
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (19)..(2037)

<400> 1
```
tgggagatcc gtgtaaag atg gga gag aga gtt cgc atc aaa ttt ggt gac      51
                    Met Gly Glu Arg Val Arg Ile Lys Phe Gly Asp
                      1               5                  10

ttt gac att gaa gat tct gat tct tgt cac ttt aat tac ttg aga att      99
Phe Asp Ile Glu Asp Ser Asp Ser Cys His Phe Asn Tyr Leu Arg Ile
             15                  20                  25

tat aat gga att gga gtc agc aga act gaa ata ggc aaa tac tgt ggt     147
Tyr Asn Gly Ile Gly Val Ser Arg Thr Glu Ile Gly Lys Tyr Cys Gly
         30                  35                  40

ctg ggg ttg caa atg aac cat tca att gaa tca aaa ggc aat gaa atc     195
Leu Gly Leu Gln Met Asn His Ser Ile Glu Ser Lys Gly Asn Glu Ile
     45                  50                  55

aca ttg ctg ttc atg agt gga atc cat gtt tct gga cgc gga ttt ttg     243
Thr Leu Leu Phe Met Ser Gly Ile His Val Ser Gly Arg Gly Phe Leu
 60                  65                  70                  75

gcc tca tac tct gtt ata gat aaa caa gat cta att act tgt ttg gac     291
Ala Ser Tyr Ser Val Ile Asp Lys Gln Asp Leu Ile Thr Cys Leu Asp
                 80                  85                  90

act gca tcc aat ttt ttg gaa cct gag ttc agt aag tac tgc cca gct     339
Thr Ala Ser Asn Phe Leu Glu Pro Glu Phe Ser Lys Tyr Cys Pro Ala
             95                 100                 105

ggt tgt ctg ctt cct ttt gct gag ata tct gga aca att cct cat gga     387
Gly Cys Leu Leu Pro Phe Ala Glu Ile Ser Gly Thr Ile Pro His Gly
        110                 115                 120

tat aga gat tcc tcg cca ttg tgc atg gct ggt gtg cat gca gga gta     435
Tyr Arg Asp Ser Ser Pro Leu Cys Met Ala Gly Val His Ala Gly Val
    125                 130                 135

gtg tca aac acg ttg ggc ggc caa atc agt gtt gta att agt aaa ggt     483
```

Val Ser Asn Thr Leu Gly Gly Gln Ile Ser Val Val Ile Ser Lys Gly
140 145 150 155

atc ccc tat tat gaa agt tct ttg gct aac aac gtc aca tct gtg gtg 531
Ile Pro Tyr Tyr Glu Ser Ser Leu Ala Asn Asn Val Thr Ser Val Val
160 165 170

gga cac tta tct aca agt ctt ttt aca ttt aag aca agt gga tgt tat 579
Gly His Leu Ser Thr Ser Leu Phe Thr Phe Lys Thr Ser Gly Cys Tyr
175 180 185

gga aca ctg ggg atg gag tct ggt gtg atc gcg gat cct caa ata aca 627
Gly Thr Leu Gly Met Glu Ser Gly Val Ile Ala Asp Pro Gln Ile Thr
190 195 200

gca tca tct gtg ctg gag tgg act gac cac aca ggg caa gag aac agt 675
Ala Ser Ser Val Leu Glu Trp Thr Asp His Thr Gly Gln Glu Asn Ser
205 210 215

tgg aaa ccc aaa aaa gcc agg ctg aaa aaa cct gga ccg cct tgg gct 723
Trp Lys Pro Lys Lys Ala Arg Leu Lys Lys Pro Gly Pro Pro Trp Ala
220 225 230 235

gct ttt gcc act gat gaa tac cag tgg tta caa ata gat ttg aat aag 771
Ala Phe Ala Thr Asp Glu Tyr Gln Trp Leu Gln Ile Asp Leu Asn Lys
240 245 250

gaa aag aaa ata aca ggc att ata acc act gga tcc acc atg gtg gag 819
Glu Lys Lys Ile Thr Gly Ile Ile Thr Thr Gly Ser Thr Met Val Glu
255 260 265

cac aat tac tat gtg tct gcc tac aga atc ctg tac agt gat gat ggg 867
His Asn Tyr Tyr Val Ser Ala Tyr Arg Ile Leu Tyr Ser Asp Asp Gly
270 275 280

cag aaa tgg act gtg tac aga gag cct ggt gtg gag caa gat aag ata 915
Gln Lys Trp Thr Val Tyr Arg Glu Pro Gly Val Glu Gln Asp Lys Ile
285 290 295

ttt caa gga aac aaa gat tat cac cag gat gtg cgt aat aac ttt ttg 963
Phe Gln Gly Asn Lys Asp Tyr His Gln Asp Val Arg Asn Asn Phe Leu
300 305 310 315

cca cca att att gca cgt ttt att aga gtg aat cct acc caa tgg cag 1011
Pro Pro Ile Ile Ala Arg Phe Ile Arg Val Asn Pro Thr Gln Trp Gln
320 325 330

cag aaa att gcc atg aaa atg gag ctg ctc gga tgt cag ttt att cct 1059
Gln Lys Ile Ala Met Lys Met Glu Leu Leu Gly Cys Gln Phe Ile Pro
335 340 345

aaa ggt cgt cct cca aaa ctt act caa cct cca cct cct cgg aac agc 1107
Lys Gly Arg Pro Pro Lys Leu Thr Gln Pro Pro Pro Pro Arg Asn Ser
350 355 360

aat gac ctc aaa aac act aca gcc cct cca aaa ata gcc aaa ggt cgt 1155
Asn Asp Leu Lys Asn Thr Thr Ala Pro Pro Lys Ile Ala Lys Gly Arg
365 370 375

gcc cca aaa ttt acg caa cca cta caa cct cgc agt agc aat gaa ttt 1203
Ala Pro Lys Phe Thr Gln Pro Leu Gln Pro Arg Ser Ser Asn Glu Phe
380 385 390 395

cct gca cag aca gaa caa aca act gcc agt cct gat atc aga aat act 1251
Pro Ala Gln Thr Glu Gln Thr Thr Ala Ser Pro Asp Ile Arg Asn Thr
400 405 410

acc gta act cca aat gta acc aaa gat gta gcg ctg gct gca gtt ctt 1299
Thr Val Thr Pro Asn Val Thr Lys Asp Val Ala Leu Ala Ala Val Leu
415 420 425

gtc cct gtg ctg gtc atg gtc ctc act act ctc att ctc ata tta gtg 1347
Val Pro Val Leu Val Met Val Leu Thr Thr Leu Ile Leu Ile Leu Val
430 435 440

tgt gct tgg cac tgg aga aac aga aag aaa aaa act gaa ggc acc tat 1395
Cys Ala Trp His Trp Arg Asn Arg Lys Lys Lys Thr Glu Gly Thr Tyr
445 450 455

gac tta cct tac tgg gac cgg gca ggt tgg tgg aaa gga atg aag cag 1443
Asp Leu Pro Tyr Trp Asp Arg Ala Gly Trp Trp Lys Gly Met Lys Gln
460 465 470 475

ttt ctt cct gca aaa gca gtg gac cat gag gaa acc cca gtt cgc tat 1491
Phe Leu Pro Ala Lys Ala Val Asp His Glu Glu Thr Pro Val Arg Tyr
480 485 490

agc agc agc gaa gtt aat cac ctg agt cca aga gaa gtc acc aca gtg 1539
Ser Ser Ser Glu Val Asn His Leu Ser Pro Arg Glu Val Thr Thr Val
495 500 505

ctg cag gct gac tct gca gag tat gct cag cca ctg gta gga gga att 1587
Leu Gln Ala Asp Ser Ala Glu Tyr Ala Gln Pro Leu Val Gly Gly Ile
510 515 520

gtt ggt aca ctt cat caa aga tct acc ttt aaa cca gaa gaa gga aaa 1635
Val Gly Thr Leu His Gln Arg Ser Thr Phe Lys Pro Glu Glu Gly Lys
525 530 535

gaa gca ggc tat gca gac cta gat cct tac aac tca cca ggg cag gaa 1683
Glu Ala Gly Tyr Ala Asp Leu Asp Pro Tyr Asn Ser Pro Gly Gln Glu
540 545 550 555

gtt tat cat gcc tat gct gaa cca ctc cca att acg ggg cct gag tat 1731
Val Tyr His Ala Tyr Ala Glu Pro Leu Pro Ile Thr Gly Pro Glu Tyr
560 565 570

gca acc cca atc atc atg gac atg tca ggg cac ccc aca act tca gtt 1779
Ala Thr Pro Ile Ile Met Asp Met Ser Gly His Pro Thr Thr Ser Val
575 580 585

ggt cag ccc tcc aca tcc act ttc aag gct acg ggg aac caa cct ccc 1827
Gly Gln Pro Ser Thr Ser Thr Phe Lys Ala Thr Gly Asn Gln Pro Pro
590 595 600

cca cta gtg gga act tac aat aca ctt ctc tcc agg act gac agc tgc 1875
Pro Leu Val Gly Thr Tyr Asn Thr Leu Leu Ser Arg Thr Asp Ser Cys
605 610 615

tcc tca gcc cag gcc cag tat gat acc ccg aaa gct ggg aag cca ggt 1923
Ser Ser Ala Gln Ala Gln Tyr Asp Thr Pro Lys Ala Gly Lys Pro Gly
620 625 630 635

cta cct gcc cca gac gaa ttg gtg tac cag gtg cca cag agc aca caa 1971
Leu Pro Ala Pro Asp Glu Leu Val Tyr Gln Val Pro Gln Ser Thr Gln
640 645 650

```
gaa gta tca gga gca gga agg gat ggg gaa tgt gat gtt ttt aaa gaa      2019
Glu Val Ser Gly Ala Gly Arg Asp Gly Glu Cys Asp Val Phe Lys Glu
            655                 660                 665

atc ctt tga aga tga tgc tgcttttta                                    2046
Ile Leu     Arg     Cys
        670


&lt;210&gt; 2
&lt;211&gt; 669
&lt;212&gt; PRT
&lt;213&gt; Homo sapiens

&lt;400&gt; 2
Met Gly Glu Arg Val Arg Ile Lys Phe Gly Asp Phe Asp Ile Glu Asp
 1               5                   10                  15

Ser Asp Ser Cys His Phe Asn Tyr Leu Arg Ile Tyr Asn Gly Ile Gly
            20                  25                  30

Val Ser Arg Thr Glu Ile Gly Lys Tyr Cys Gly Leu Gly Leu Gln Met
        35                  40                  45

Asn His Ser Ile Glu Ser Lys Gly Asn Glu Ile Thr Leu Leu Phe Met
    50                  55                  60

Ser Gly Ile His Val Ser Gly Arg Gly Phe Leu Ala Ser Tyr Ser Val
65                  70                  75                  80

Ile Asp Lys Gln Asp Leu Ile Thr Cys Leu Asp Thr Ala Ser Asn Phe
                85                  90                  95

Leu Glu Pro Glu Phe Ser Lys Tyr Cys Pro Ala Gly Cys Leu Leu Pro
            100                 105                 110

Phe Ala Glu Ile Ser Gly Thr Ile Pro His Gly Tyr Arg Asp Ser Ser
        115                 120                 125

Pro Leu Cys Met Ala Gly Val His Ala Gly Val Val Ser Asn Thr Leu
    130                 135                 140

Gly Gly Gln Ile Ser Val Val Ile Ser Lys Gly Ile Pro Tyr Tyr Glu
145                 150                 155                 160

Ser Ser Leu Ala Asn Asn Val Thr Ser Val Val Gly His Leu Ser Thr
                165                 170                 175

Ser Leu Phe Thr Phe Lys Thr Ser Gly Cys Tyr Gly Thr Leu Gly Met
            180                 185                 190

Glu Ser Gly Val Ile Ala Asp Pro Gln Ile Thr Ala Ser Ser Val Leu
        195                 200                 205

Glu Trp Thr Asp His Thr Gly Gln Glu Asn Ser Trp Lys Pro Lys Lys
    210                 215                 220

Ala Arg Leu Lys Lys Pro Gly Pro Pro Trp Ala Ala Phe Ala Thr Asp
225                 230                 235                 240

Glu Tyr Gln Trp Leu Gln Ile Asp Leu Asn Lys Glu Lys Lys Ile Thr
                245                 250                 255
```

```
Gly Ile Ile Thr Thr Gly Ser Thr Met Val Glu His Asn Tyr Tyr Val
            260                 265                 270

Ser Ala Tyr Arg Ile Leu Tyr Ser Asp Asp Gly Gln Lys Trp Thr Val
        275                 280                 285

Tyr Arg Glu Pro Gly Val Glu Gln Asp Lys Ile Phe Gln Gly Asn Lys
    290                 295                 300

Asp Tyr His Gln Asp Val Arg Asn Asn Phe Leu Pro Pro Ile Ile Ala
305                 310                 315                     320

Arg Phe Ile Arg Val Asn Pro Thr Gln Trp Gln Gln Lys Ile Ala Met
                325                 330                 335

Lys Met Glu Leu Leu Gly Cys Gln Phe Ile Pro Lys Gly Arg Pro Pro
            340                 345                 350

Lys Leu Thr Gln Pro Pro Pro Pro Arg Asn Ser Asn Asp Leu Lys Asn
        355                 360                 365

Thr Thr Ala Pro Pro Lys Ile Ala Lys Gly Arg Ala Pro Lys Phe Thr
    370                 375                 380

Gln Pro Leu Gln Pro Arg Ser Ser Asn Glu Phe Pro Ala Gln Thr Glu
385                 390                 395                     400

Gln Thr Thr Ala Ser Pro Asp Ile Arg Asn Thr Thr Val Thr Pro Asn
                405                 410                 415

Val Thr Lys Asp Val Ala Leu Ala Ala Val Leu Val Pro Val Leu Val
            420                 425                 430

Met Val Leu Thr Thr Leu Ile Leu Ile Leu Val Cys Ala Trp His Trp
        435                 440                 445

Arg Asn Arg Lys Lys Lys Thr Glu Gly Thr Tyr Asp Leu Pro Tyr Trp
    450                 455                 460

Asp Arg Ala Gly Trp Trp Lys Gly Met Lys Gln Phe Leu Pro Ala Lys
465                 470                 475                     480

Ala Val Asp His Glu Glu Thr Pro Val Arg Tyr Ser Ser Ser Glu Val
                485                 490                 495

Asn His Leu Ser Pro Arg Glu Val Thr Thr Val Leu Gln Ala Asp Ser
            500                 505                 510

Ala Glu Tyr Ala Gln Pro Leu Val Gly Gly Ile Val Gly Thr Leu His
        515                 520                 525

Gln Arg Ser Thr Phe Lys Pro Glu Glu Gly Lys Glu Ala Gly Tyr Ala
    530                 535                 540

Asp Leu Asp Pro Tyr Asn Ser Pro Gly Gln Glu Val Tyr His Ala Tyr
545                 550                 555                     560

Ala Glu Pro Leu Pro Ile Thr Gly Pro Glu Tyr Ala Thr Pro Ile Ile
                565                 570                 575

Met Asp Met Ser Gly His Pro Thr Thr Ser Val Gly Gln Pro Ser Thr
            580                 585                 590
```

Ser Thr Phe Lys Ala Thr Gly Asn Gln Pro Pro Pro Leu Val Gly Thr
595 600 605

Tyr Asn Thr Leu Leu Ser Arg Thr Asp Ser Cys Ser Ser Ala Gln Ala
610 615 620

Gln Tyr Asp Thr Pro Lys Ala Gly Lys Pro Gly Leu Pro Ala Pro Asp
625 630 635 640

Glu Leu Val Tyr Gln Val Pro Gln Ser Thr Gln Glu Val Ser Gly Ala
645 650 655

Gly Arg Asp Gly Glu Cys Asp Val Phe Lys Glu Ile Leu
660 665

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US99/19047

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) :C07H 21/04; C07K 1/00; C12N 5/00
US CL : 536/23.2; 530/350; 800/8
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 536/23.2; 530/350; 800/8

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

STN Registry File- SEQ ID NO:1 and 2

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	GWO-SHE LEE, M. et al. Evolutionary History of a Multigene Family: An Expressed Human Beta-Tubulin Gene and Three Processed Pseudogenes. Cell. June 1983, Volume 33, pages 477-487, see entire document.	1, 2

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
27 OCTOBER 1999	29 NOV 1999

Name and mailing address of the ISA/US	Authorized officer
Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231	BRADLEY L. SISSON D. Lawrence for
Facsimile No. (703) 305-3230	Telephone No. (703) 308-0196

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US99/19047

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☒ Claims Nos.: 14
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

Claim 14 was not drawn to any specific invention but rather was drawn to "any invention described" within the application. It is not clear what applicant contempaltes the invention being and as such a search of same cannot be conducted.

3. ☒ Claims Nos.: 5, 8-10, 12 and 13
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest
☐ The additional search fees were accompanied by the applicant's protest.
☐ No protest accompanied the payment of additional search fees.